

뉴로-퍼지 기법에 의한 오존농도 예측모델 Neuro-Fuzzy Approaches to Ozone Prediction System

김태헌 · 김성신 · 김인택* · 이종범** · 김신도*** · 김용국**

Tae-Hun Kim, Sungshin Kim, In-Taek Kim*, Chong-Bum Lee**
Shin-Do Kim*** and Yong-Goog Kim**

부산대학교 전기공학과, * 명지대학교 전기정보공학부
** 강원대학교 환경학과, *** 서울시립대학교 환경공학과

ABSTRACT

In this paper, we present the modeling of the ozone prediction system using Neuro-Fuzzy approaches. The mechanism of ozone concentration is highly complex, nonlinear, and nonstationary, the modeling of ozone prediction system has many problems and the results of prediction is not a good performance so far. The Dynamic Polynomial Neural Network(DPNN) which employs a typical algorithm of GMDH(Group Method of Data Handling) is a useful method for data analysis, identification of nonlinear complex system, and prediction of a dynamical system. The structure of the final model is compact and the computational speed to produce an output is faster than other modeling methods. In addition to DPNN, this paper also includes a Fuzzy Logic Method for modeling of ozone prediction system. The results of each modeling method and the performance of ozone prediction are presented. The proposed method shows that the prediction to the ozone concentration based upon Neuro-Fuzzy approaches gives us a good performance for ozone prediction in high and low ozone concentration with the ability of superior data approximation and self organization.

1. 서 론

최근 국내외를 막론하고 하계 중 대류권의 고농도 오존의 발생은 대기오염에 의해 발생하는 현상 중 하나로 많은 관심을 끌고 있다. 일반적으로 오존은 질소산화물 및 탄화수소류등의 오염 물질로 인한 광화학적인 작용과 일사량, 풍속, 기온 등의 기상학적인 변수들의 상호작용으로 생성되며, 지구에 존재하는 오존의 약 10%정도가 지표에서 10 km이내인 대류권에 존재하고 있다. 오존은 주로 하계인 6월부터 8월 사이에 집중적인 고농도 현상을 보이고 있으며, 강력한 산화력으로 인해 적당량의 오존은 인간에게 이로우나 기준농도 이상의 오존은 인체의 건강이나 농작물의 수확 및 생태계에 피해를 주고 있다. 특히 고농도 오존이 발생하는 날은 호흡기 질환을 가진 노약자들의 사망이 증가했다는 통계를 얻을 수 있다. 따라서 이러한 오존의 피해를 줄이기 위해 대기오염 예보제 실시와 함께 기상학적 및 광화학적 특성에 의한 오존의 농도 변화 및 정확한 고농도 오존예측에 관한 연구가 진행되고 있다. 기존의 고농도 오존예측을 위한 방법으로 통계적 방법에 의한 선형 회귀 모델[9], 다변량 통계분석[7]에 의한 방법, 그리고 신경회로망 등의 인공지능 기법을 이용한 예측모델[2]에 의한 연구가 이루어져 왔다. 그러나 기상학적인 변수와 오존 오염물질간의 강한 비선형성과 대류권내에서의 오존생성에 관한

매우 복잡한 반응작용으로 인하여 고농도 오존을 예측할 때는 대부분의 결과가 실제 오존농도보다 낮게 예측하는 경향을 보이고 있다. 그 이유로는 기존의 오존 농도 예측 모델링이 비선형적이고, 비정상적인(nonstationary) 경향의 적절한 반영이 이루어지지 않았거나, 예측에 절대적 영향을 미치는 고농도 오존 데이터의 부족, 모델링에 필요한 입력요소를 적절히 선택하지 못한 경우로 나눌 수 있다.

본 논문에서는 보다 정확한 모델을 얻기 위해 입력과 출력사이에 존재하는 비선형성을 적극 반영하고 모델링에 필요한 입력 요소를 적절하게 선택하기 위해 동적인 다항식 신경회로망(Dynamic Polynomial Neural Network)[1]과 퍼지 논리시스템(Fuzzy Logic System)을 적용한다. 신경회로망과 퍼지 시스템은 잘 알려진 바와 같이 비선형적인 학습이 가능하므로 비선형적인 입·출력간의 관계를 잘 표현할 수 있다. 한편 예측기의 입력은 매우 다양한 요소가 존재하는데, 이들 중 출력, 즉 오존 농도의 예측치와 밀접한 관계의 요소를 찾기 위해서는 동적인 다항식 신경회로망을 사용한다.

본 논문은 다음과 같이 구성되어 있다. 2절에서는 동적인 다항식 신경회로망에 대해 설명한다. 여기에는 동적 다항식 신경회로망의 일반적인 구조, 예측모델, 모델구조에 대해서 언급한다. 3절은 퍼지 예측기로 일반적인 퍼지논리 시스템과 퍼지 예측기의 구조를 설

명한다. 신경회로망과 퍼지 예측기를 적용한 예측 결과는 4절에서 다룬다. 학습과 예측에는 동일한 장소의 데이터를 이용해야 하므로 서울 쌍문동에서 얻은 측정치를 사용하였다. 마지막으로 5절은 결론 부분으로 두 예측기의 예측 성능에 대한 간단한 비교와 앞으로 수행방향을 요약하였다.

2. 동적 다항식 신경회로망

2.1 동적 다항식 신경회로망의 기본구조

다량의 관측자료와 변수들로부터 시스템의 모델을 구성하기 위해 GMDH(Group Method of Data Handling)[1,4,11]를 이용한 동적 다항식 신경회로망은 비선형적이고 복잡한 동적인 시스템의 모델링과 예측 및 지능제어에 응용되어지고 있다.

각 노드에 대해서 두 개의 입력변수로부터 하나의 출력을 생성하는 기본적인 형태의 동적 다항식 신경회로망의 구조를 그림 1에 나타내었으며, 동적 다항식 신경회로망의 각 노드에서 입력변수와 출력과의 관계는 다음과 같은 함수식으로 표현된다.

$$\begin{aligned} y_1 &= \omega_{01} + \omega_{11}x_1 + \omega_{21}x_2 + \omega_{31}x_1x_2 + \omega_{41}x_1^2 + \omega_{51}x_2^2 \\ y_2 &= \omega_{02} + \omega_{12}x_3 + \omega_{22}x_4 + \omega_{32}x_3x_4 + \omega_{42}x_3^2 + \omega_{52}x_4^2 \end{aligned} \quad (1)$$

각 노드에서의 출력 y_1, y_2 는 아래 식과 같이 입력 변수들의 결합에 의한 다항식으로 나타내어지며, 최종 출력 z 는 각 노드의 출력 값 y_1, y_2 를 입력변수로 사용하여 다른 형태의 다항식으로 나타내어진다.

$$z = \omega_{03} + \omega_{13}y_1 + \omega_{23}y_2 + \omega_{33}y_1y_2 + \omega_{43}y_1^2 + \omega_{53}y_2^2 \quad (2)$$

만약 각 노드에서 입력의 변수가 세 개일 경우에는 새로운 입력 조합항과 변수들의 3차 항이 추가된다.

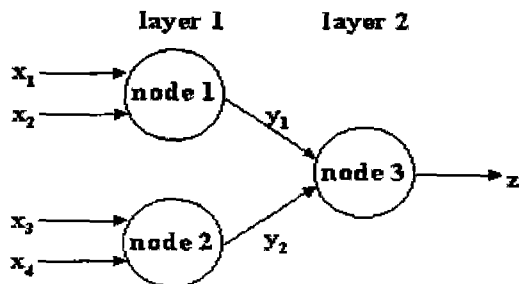


그림 1. 동적 다항식 신경회로망의 기본구조
Fig. 1. The basic structure of Dynamic Polynomial Neural Network

각 노드의 파라미터 $w_{ij}(i=0, 1, 2, \dots, n, j=0, 1, 2, \dots, k)$ 를 결정하기 위해 Least Square Method를 사용하며, 목적함수인 실제 측정된 값과 학습된 출력 값과의 오차를 최소화하는 파라미터 w_j 를 학습 데이터를 이용하여 구하게 된다.

식 (3),(4)는 목적함수와 파라미터 w_j 의 계산식을 보여주고 있다. 이렇게 해서 구해진 파라미터가 현재 노드의 출력함수를 구성하고, 다음 층 노드의 입력으로 들어가게 되며 이러한 과정을 반복하여 최종적으로 우수한 출력 값을 나타내는 함수를 얻을 수 있다.

$$J = \sum_{k=1}^{\# \text{ of data}} (z(k) - \hat{z}(k))^2 = \|z - \Phi w\|^2 \quad (3)$$

$$w = (\Phi^T \Phi)^{-1} \Phi^T z \quad (4)$$

동적 다항식 신경회로망을 이용한 모델링 방법의 장점은 연속적인 기본구조로 구성되어 있으며, 많은 입력변수를 사용하여 마지막 출력의 계산까지 빠른 시간동안에 종결할 수 있다. 또한 사용 가능한 입력변수들을 사용한 후 최종 모델의 구조에 따라서 입력변수가 결정되어 중요한 변수를 선택할 수 있다는 장점이 있다. 처음에 사용된 입력변수들은 모두가 결과에 영향을 미치는 것이 아니라 선택 방법에 따라서 결과에 영향을 미치는 주요 입력변수들이 결정되어진다. 최종적인 모델의 출력은 모든 입력성분에 의하여 영향을 받는 것이 아니라, 선택되어진 입력에 의하여 출력이 결정되기 때문에 주요 입력변수를 선택할 수 있는 장점이 있다.

2.2 모델의 성능 평가 및 선택

동적 다항식 신경회로망의 또 다른 특징으로서 자기 조직화(Self Organization)하는 능력을 들 수 있는데, 동적 다항식 신경회로망은 GMDH방법을 사용하여 입력 자료를 분산에 따라 학습 데이터와 테스트 데이터로 나누어서 시스템의 모델링 과정 중에 나타날 수 있는 학습데이터로의 편중현상을 피하도록 한다.

두 개로 나누어진 데이터 집단을 이용하여 모델의 구조와 각 노드에서의 파라미터를 구하게 되는데, 일반적으로 학습 데이터를 이용하여 파라미터를 구하고, 테스트 데이터를 이용하여 모델의 성능을 평가하며, 학습 오차와 테스트 오차를 이용하여 모델의 최종적인 구조를 결정하게 된다. 따라서 동적 다항식 신경회로망은 각 층에서 선택기준(Selection Criterion)에 의해 다음 층의 입력을 선택하고, 그림 2에 나타낸 것처럼 학습 오차와 테스트 오차를 이용한 PC

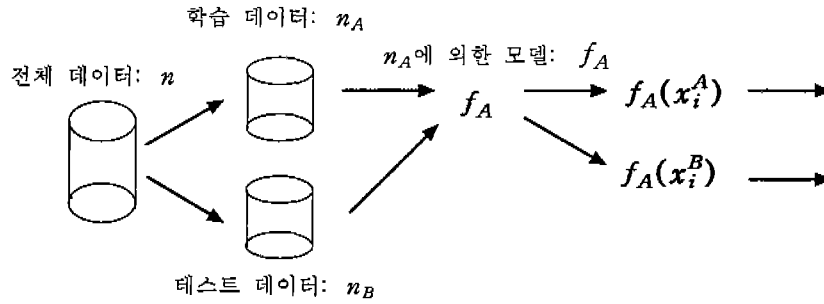


그림 2. 모델선택을 위한 데이터 분리
 Fig. 2. Data split for model selection

(Performance Criterion)에 따라서 모델의 입력변수와 구조를 스스로 결정한다.

식 (5)에 학습 오차와 테스트 오차를 이용하여 PC 를 결정하기 위한 식을 나타내었으며, 가중치 η 는 0에서 1사이의 값을 사용하였다. 식 (5)에 의한 성능평가 방법은 구성된 모델이 학습데이터 뿐만 아니라 테스트 데이터에도 잘 적용하는가를 평가하며 준비되지 않은 새로운 데이터에도 적용되는 모델평가방법이다.

$$e_1^2 = \sum_{i=1}^{n_A} (y_i^A - f_A(x_i^A))^2 / n_A,$$

$$e_2^2 = \sum_{i=1}^{n_B} (y_i^B - f_A(x_i^B))^2 / n_B,$$

$$PC = e_1^2 + e_2^2 + \eta(e_1^2 - e_2^2)^2, \quad (5)$$

f_A 는 학습 데이터로부터 구성된 모델이며, e_1 는 학습오차, e_2 는 테스트 오차, n_A 는 학습 데이터의 개수, n_B 는 테스트 데이터의 개수, y_i 는 실제 측정된 값이며, $f_A(x_i^A)$ 와 $f_B(x_i^B)$ 는 각각 학습에 의한 결과 값과 테스트에 의한 결과 값이다. 전체 데이터의 개수는 $n = n_A + n_B$ 이다. 따라서 최적의 모델은 학습 오차와 테스트 오차에 대해 PC 값이 최소가 되는 곳에서 얻어진다.

η 는 e_1 과 e_2 의 차이에 대한 가중치이다. 우리는 n_A 와 n_B 에 대해 최고의 모델을 선택하기 위해 PC 를 최소화하는 것을 원한다. e_1 과 e_2 에 대해 PC 를 편미분하면 아래와 같다.

$$\frac{\partial PC}{\partial e_1} = 4\eta e_1^3 + 2(1 - 2\eta e_2^2)e_1 \quad (6)$$

$$\frac{\partial PC}{\partial e_2} = 4\eta e_2^3 + 2(1 - 2\eta e_1^2)e_2 \quad (7)$$

식 (7)은 e_1 에 대해 PC 를 편미분한 것은 e_1 과 e_2 의 함수임을 나타낸다. 또한, 식 (7)과 (8)은 e_1 과 e_2 에

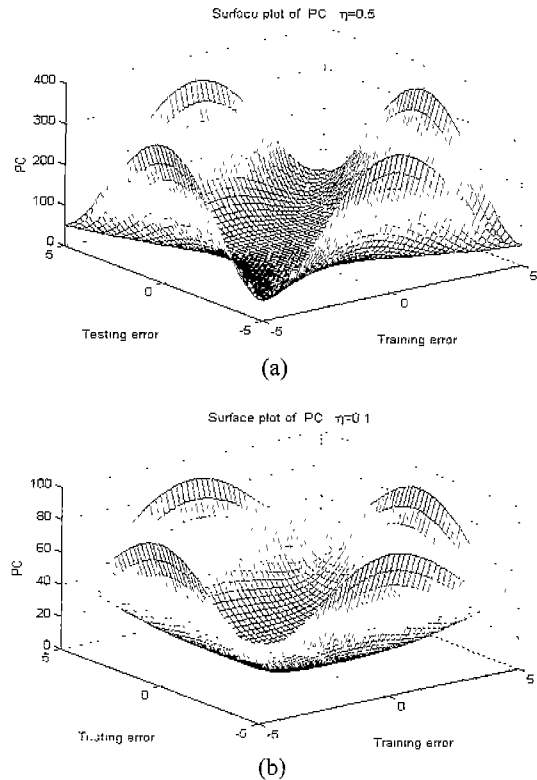


그림 3. η 값에 따른 PC 에 대한 3차원 (a) $\eta = 0.5$ 일 때, (b) $\eta = 0.1$ 일 때

Fig. 3. Surface plot of PC (a) Surface plot with $\eta = 0.5$, (b) Surface plot with $\eta = 0.1$

대하여 대칭이다. PC 에 의한 최종 모델은 학습데이터에만 최적화 되지 않고 학습과 테스트 데이터 모두를 만족시켜야 한다. 따라서 최적의 모델은 학습 오차와 테스트 오차에 대해 PC 가 최소화되는 곳에서 얻어진다. 그림 3에 η 의 값의 변화에 따른 PC 값의 변화를 나타내었다.

PC 의 변화를 살펴보기 위한 예제로서, 그림 4는 등

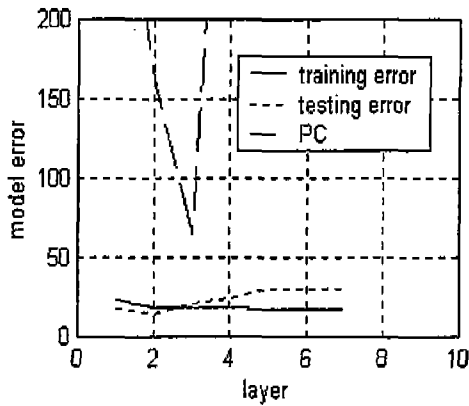


그림 4. 층수의 증가에 따른 모델성능의 변화
Fig. 4. Variation of PC according to the change of layer

적 다항식 신경회로망의 층수의 증가에 따른 학습 오차와 테스트 오차, 그리고 PC 값의 변화를 나타냈으며, PC 값이 3층에서 최소가 됨을 알 수 있다.

2.3 예측 모델

대기오염물질의 시간에 따른 변화량 예측은 관측된 변수들과 그 변수들로 구성된 예보 시스템 모델의 정확성에 의하여 결정된다. 예측모델의 구성을 위하여 물리·화학적 현상의 수학적 표현을 이용할 수 있으나, 대기의 현상은 매우 복잡한 동적인 시스템이므로, 본 연구에서는 대기현상과 오염물질에 대한 축적된 데이터로부터 동적인 동적 다항식 신경회로망을 이용하여 예측모델을 구성하고 그 성능을 향상시키고자 한다.

일반적으로 오존의 농도예측을 위해 사용되는 대기 오염자료로서는 오존, 일산화탄소, 이산화질소, 아황산가스, 부유 먼지 등이 사용되며, 기상자료로서 풍속, 풍향, 기온, 일사량, 강수량, 습도 및 운량 등이 사용

표 1. 모델링에 사용된 변수들
Table 1. Variables for ozone modeling

구분	변수	시간	데이터 범위	
입력 변수	오존(o3)	06시~09시	0~57ppb/h	
	오염 자료	이산화질소(NO2)	06시~09시	7~105ppb/h
		일산화탄소(Co)	06시~09시	0.1~3.1ppm/h
		기온(tair)	14시~17시	13.7~35.7°C/h
	기상 자료	상대습도(th)	14시~17시	14~94g/m ³ /h
		일사량(sr)	14시~17시	1~325MJ/m ² /h
풍속(ws)		14시~17시	0~7.7m/sec ² /h	
출력 변수	오존(o3)	14시~17시	2~162ppb/h	

표 2. 학습에 사용된 자료
Table 2. Data for training process

입력변수							출력변수
오존	질소	탄소	기온	상대습도	일사량	풍속	14시오존
38	37	8	147	82	17	42	28
34	40	8	156	77	67	40	30
46	23	5	161	76	39	22	33
34	36	6	164	73	55	27	37
37	24	6	199	22	138	43	49
34	29	6	205	23	296	55	51
35	22	7	207	20	204	33	53
40	21	5	214	17	273	45	54
29	57	20	234	17	177	8	44
34	62	20	248	14	229	2	46
45	82	28	238	15	46	10	48
50	77	29	243	15	171	10	50
27	41	6	147	75	63	23	31
27	48	7	151	70	57	15	31
29	46	7	155	66	117	12	36
21	64	10	158	61	100	8	38
34	24	6	199	17	263	27	36
33	24	6	192	17	299	52	37
27	33	7	195	19	212	38	37
31	31	6	192	22	145	38	38
17	52	13	187	28	97	55	35
19	55	13	200	25	155	25	40
12	51	9	205	25	242	23	41
15	56	11	214	22	175	23	42
21	60	13	204	43	31	47	30
13	53	10	220	35	178	18	37
17	56	13	218	38	63	37	43
15	49	10	232	33	195	18	43
24	38	9	174	41	149	53	29
25	34	8	179	40	218	42	34
27	37	8	185	39	245	53	36
27	34	7	180	42	280	40	37
20	51	11	179	41	155	50	39
18	50	15	183	40	234	48	43
14	50	12	187	38	255	45	44
12	57	12	185	37	263	50	45
10	50	10	203	34	254	23	44
15	51	10	212	30	239	18	45
18	52	10	213	31	193	22	47
21	59	12	213	31	85	23	48
36	21	6	189	40	284	62	46
33	25	6	190	39	270	40	48
35	29	7	170	42	122	35	49
34	28	8	180	41	218	37	51
17	41	12	195	41	296	40	47
20	58	14	195	38	161	43	48
15	41	12	202	37	274	40	50
17	50	13	199	37	226	47	51
15	57	15	207	38	117	53	45
25	65	18	208	37	158	40	46
23	65	14	208	35	99	37	46
19	57	15	209	38	123	28	47

된다. 이러한 자료 중에서 100 ppb이상의 고농도 오존에서의 강수량은 대부분 0mm이기 때문에 고농도 오존예측에 거의 영향을 미치지 못하며, 또한 풍향과 운량은 수치화의 어려움, 그리고 아황산가스와 부유 먼지는 대기 오염규제로 인해 그 배출량이 줄어들고 있기 때문에 입력자료에서 삭제하였다. 따라서 학습될 변수로 오존, 일산화탄소, 이산화질소, 기온, 일사량, 상대습도, 풍속 등을 사용하였다. 표 1에 오존 농도예측에 사용된 변수들을 나타내었으며, 쌍문동에 대한 자료의 리스트를 표 2에 나타내었다. 표 2에서 기온, 일산화탄소, 풍속은 실제 값에 10을 곱한 수치이다.

오존 농도는 하루 중 14~17시경에 최고농도를 나타내기 때문에 이 시간대의 오존 농도를 예측하는 것을 목표로 하였으며, 오염 자료중 이산화질소, 오존, 일산화탄소는 오전의 자료를 사용하고 나머지 자료는 오존 농도 예측 시간대의 자료를 사용하였다.

서울지역 중 하계에 고농도 오존이 자주 발생하는 지역인 쌍문동과 방이동, 불광동 지역의 96년과 97년도 5월 달부터 9월 달 사이의 자료를 이용하였고, 결국된 데이터를 위하여 Spline 보간법을 사용하였다. 그림 5에 97년도 쌍문동 지역의 하계 중 100 ppb 이상의 오존농도가 발생한 날과 오존 농도의 변화를 나타내었다.

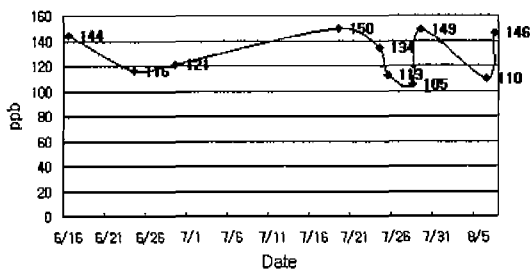


그림 5. 97년도 쌍문동 오존 농도변화(100 ppb이상)
Fig. 5. Variation of high ozone concentration in Ssang-Mun-Dong, 1997

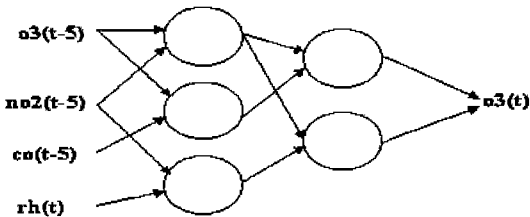


그림 6. 선택된 입력변수와 모델의 구조-1
(t:14시, 데이터의 개수:153개, o3:오존, no2:이산화질소, co:일산화탄소, rh:상대습도)

Fig. 6. Selected input variable and model structure-1

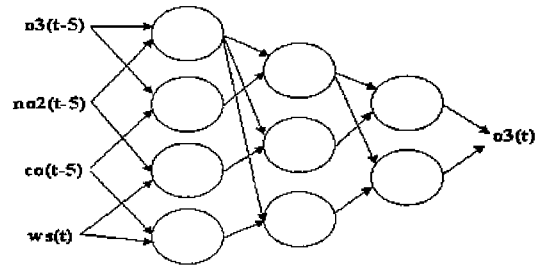


그림 7. 선택된 입력변수와 모델의 구조-2
(t:14시, 데이터의 개수:153개, o3:오존, no2:이산화질소, co:일산화탄소, ws:풍속)

Fig. 7. Selected input variable and model structure-2

2.4 모델 구조

위에서 선정한 7개의 기상자료와 오염물질자료에 대해서 14시~17시의 각 시간대별 오존의 농도를 출력으로 두고 나머지 7개 변수의 5시간전의 자료를 입력으로 하여 학습을 시킨 결과 선택된 입력변수들과 모델의 네트워크 구조를 아래 그림 6과 그림 7에 나타내었다. 각 노드에서 두 개의 입력변수의 조합으로 하나의 출력을 생성하는 구조를 사용하였으며, 학습을 위하여 사용된 데이터 수는 153개이다.

그림 6은 방이동 14시 예측모델의 구조이며, 그림 7은 쌍문동 14시 예측 모델의 구조이다. 그림 6과 그림 7에서 학습할 데이터의 차이에 따라 선택되는 입력변수와 모델의 구조가 달라짐을 알 수 있다.

3. 퍼지 예측기

3.1 퍼지논리 시스템

앞 장에서 설명한 신경회로망과 함께 1980년대 이후 연산 지능 (computational intelligence)의 분야에서 주목을 받았던 퍼지논리 시스템도 보편적인 근사기 (universal approximator)의 기능을 가지고 있다[18, 19]. 이러한 보편적인 근사기는 주어진 데이터의 학습을 통해 데이터의 재현뿐만 아니라 미래의 값을 예측할 수 있는 기능을 가지고 있다. 따라서 퍼지논리를 사용한 예측기는 비선형적인 신호뿐만 아니라, 복잡하면서도 불규칙적인 신호로 보이는 카오스 신호에 쉽게 적용할 수 있다[20]. 이러한 비선형 신호, 카오스 신호는 결정론적인 식(deterministic equation)에 의해 표현이 가능하므로 이를 퍼지규칙으로 표현하는 것은 오히려 당연하다고 볼 수 있다.

퍼지논리 시스템을 학습시키기 위한 방법에는 여러 종류가 있는데, 예로 역전파학습(back-propagation), 직교최소자승법(orthogonal least squares), 최근접 클러

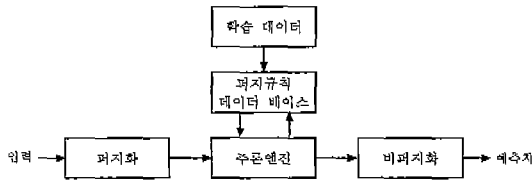


그림 8. 퍼지논리 시스템의 구성도
Fig. 8. Structure fuzzy logic system

스터링(nearest neighbor clustering), 룩업테이블(lookup table) 생성법 등을 들 수 있다[21]. 본 논문에서는 이들 중 가장 간단한 방법인 룩업테이블 생성법을 선택하여 오존 농도 예측에 적용하고자 한다. 이 방법은 입력과 출력 쌍으로부터 if-then 퍼지 규칙을 생성하게 되는데, 이때 입력과 출력의 범위를 고정시킨 상태에서 학습하게 되므로 이들의 값을 미리 알고 있어야 하는 단점이 있다. 그러나 이 방법은 알려진 다른 방법에 비해서 그 계산량이 매우 적기 때문에 실제로 구현하는 경우, 가장 쉽게 사용할 수 있는 큰 장점이 있다.

오존 농도의 예측기를 구현하는데 있어 가장 중요한 점은 학습된 상황 이외의 현상은 발생하지 않는다는 가정이다. 그러나 학습에 사용한 데이터의 양이 충분하지 않기 때문에 이러한 가정이 유효한지는 실제 데이터와 예측 데이터 사이에 가지는 예측 에러가 그 사실을 설명해 준다. 뿐만 아니라 경제 지수와 같이 시간에 따라 증가하는 비정상적인 경향을 포함할 수도 있지만, 이 사실 역시 충분하지 않은 데이터에서 발견하기가 매우 어렵다.

3.2 퍼지 예측기

퍼지 논리 시스템에 의한 오존 농도 예측기는 그림 8과 같은 단위 블록들로 구성되어 있다. 퍼지 규칙 데이터베이스는 퍼지 규칙들의 집합으로 구성되어 있으며, 각각의 규칙은 잘 알려진 아래와 같은 if-then 규칙으로 이루어져 있다.

$$\text{If } x \text{ is } A \text{ and } y \text{ is } B \text{ and } z \text{ is } C \dots, \text{ Then } w \text{ is } P. \quad (6)$$

여기서 x, y, z, \dots, w 는 실제의 데이터를 나타내는 변수이고, A, B, C, \dots, P 는 퍼지 집합을 의미한다.

오존 농도의 예측에서는 if 이하의 전건부의 변수는 오존농도, 이산화질소농도, 일산화탄소농도, 기온, 상대습도, 일사량, 풍속이며, then 이하의 후건부의 변수는 예측하고자 하는 오존농도이다. 전건부의 변수는 2개의 소속함수만을 가진다고 가정하였으며 후건부의

변수는 11개의 소속함수를 가진다. 따라서 전건부의 소속함수는 농도, 기온, 습도, 일사량, 풍속의 낮음과 높음의 2가지 경우만을 가진다.

전건부의 소속함수가 단순화시켜 적은 이유는 전건부의 변수가 많기 때문에 발생할 수 있는 출력의 경우의 수를 최소화시킬 필요가 있기 때문이다. 예를 들면 전건부의 소속함수가 두 가지인 경우에 대해 발생할 수 있는 출력의 경우의 수는 $128(2^7)$ 인데 비하여, 소속함수가 세 가지가 되면 출력의 경우의 수가 $2187(3^7)$ 이 된다. 따라서 이 때 우려되는 점은 주어진 입력에 대하여 해당하는 퍼지 규칙이 없어 예측기의 출력이 존재하지 않게 된다.

전건부의 소속함수의 수가 적은 것은 학습시 전건부는 같으나 후건부가 다른 모순된 규칙(conflicting rule)을 발생시킬 수 있다. 이를 위해서 아래와 같은 방법에 의해 규칙을 선택한다. 예를 들어 학습과정에서 아래와 같이 세 개의 입력 데이터로부터 각각 다른 출력을 얻었다고 가정하자.

$$x \text{ is } A \text{ and } y \text{ is } B \text{ and } z \text{ is } C, \text{ then } w \text{ is } P. \quad (7)$$

$$x \text{ is } A \text{ and } y \text{ is } B \text{ and } z \text{ is } C, \text{ then } w \text{ is } Q. \quad (8)$$

이러한 상황을 해결하기 위해서는 각 변수의 소속함수값들 중 최소값을 비교하여 최대값을 발생시킨 규칙을 선택한다. 구체적으로 예를 들면 두 집단의 학습데이터가 식 (7)을 $\{\mu(x \in A) = 0.7, \mu(y \in B) = 0.7, \mu(z \in C) = 0.7, \mu(w \in P) = 0.7\}$ 의 값으로 만족하고, 식 (8)은 $\{\mu(x \in A) = 0.8, \mu(y \in B) = 0.9, \mu(z \in C) = 0.7, \mu(w \in P) = 0.6\}$ 의 값을 가질 때 식 (7)이 식 (8)에 비해 그 최소값(식 (7)은 0.7, 식 (8)은 0.6)이 크므로 전건부의 변수들이 각각 A, B, C에 속할 때 출력은 P라고 판정하게 된다. 이와 같이 모순된 규칙을 해결하는 것은 실제로 학습 과정에서 자주 발생하므로 반드시 처리되어야 한다. 특히 전건부의 소속함수의 개수가 적을수록 모순된 규칙의 발생 가능성이 더욱 높아진다.

추론과정에서는 잘 알려진 Mamdani의 퍼지 추론 방법을 사용하였으며 비퍼지화를 위해서는 면적 중심(Center of Area) 방법을 사용하여 오존의 예측치를 발생하였다.

4. 예측 결과

4.1 동적 다항식 신경회로망

동적 다항식 신경회로망에 의해서 모델링된 오존농도 예측 시스템의 시뮬레이션 환경을 그림 9에 나타

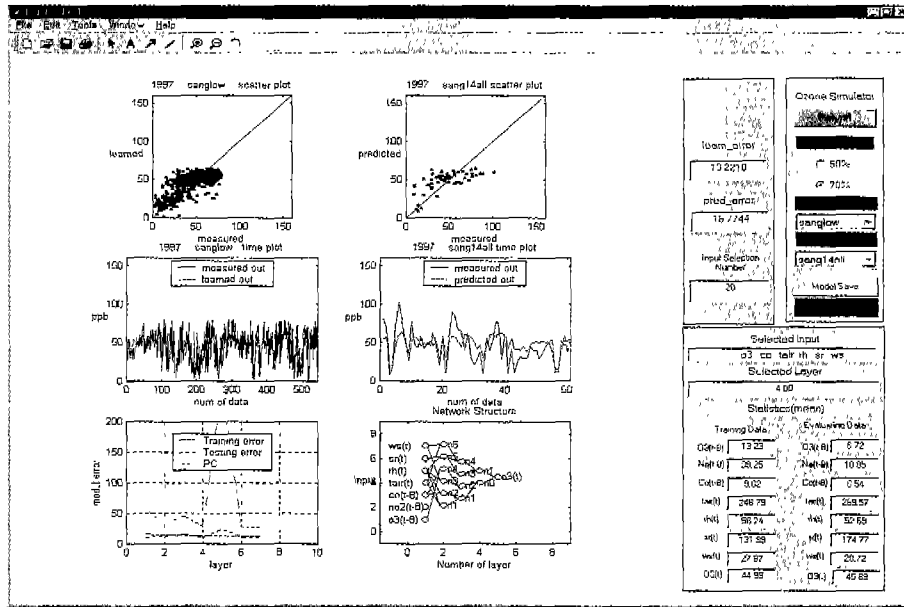


그림 9. 쌍문동 시뮬레이션 환경
Fig. 9. Environment of simulation tool for Ssang-Mun-Dong

내었다. 예측모델의 프로그램을 위하여 Matlab® (version 5.3)의 GUI(Graphic User Interface)를 사용하였으며, 쌍문동에서 실제 측정된 96년도 5월 달부터 9월 달까지의 자료와 97년 6월 달부터 7월 달까지의 자료를 학습자료로 사용하였다. 예측대상은 위의 2.3절에서 언급한 것과 같이 쌍문동 지역의 97년 8월 및 9월 달의 14시부터 17시사이의 각 시각의 오존 농도로 정하였다.

모델링 과정에서 저농도와 고농도 및 전체농도로 구분하여 모델을 구성하였고, 8시간 전의 자료를 사용하여 8시간 후의 오존 농도를 예측하는 것을 목표로 하였다.

예측을 위한 오존의 농도는 80 ppb이상과 이하, 그리고 전체 농도로 구분하였다.

데이터 분리는 학습 부분에 70%, 테스트 부분에 30%를 사용하였으며, 현재 층의 출력 결과 중 오차 (RMSE: root-mean-squared error)가 적은 20개를 다음 층의 입력으로 선택하였다. 그리고 모델의 구조가 복잡해지는 것을 방지하기 위해 층의 수는 PC의 값에 따라 3층에서 5층 사이에서 선택되도록 제한하였다.

그림 10과 11에 학습되는 데이터의 차이에 따른 모델구조의 변화를 나타내었다. 그림 10은 쌍문동 저농도 모델의 구조이며, 그림 11은 쌍문동 고농도 모델의 구조이다. 기본적인 입력변수는 2.3절에서 제시한

7개를 사용하였다. 그림 10에서는 이산화질소($\text{No}_2(t-8)$), 그림 11에서는 풍속($ws(t)$)과 오존($o_3(t-8)$)이 입력 변수로 선택되지 않았음을 볼 수 있으며, 모델의 구조는 새로운 학습데이터에 따라 새로이 구성되어진다. 따라서 모델의 구조는 고정되어있지 않고, 최적의 출력을 위하여 주어진 데이터로부터 모델의 선택과정에 따라서 변화되어진다.

표 1과 2에는 데이터 분리에 따라 선택되는 변수들이 어떻게 변하는지를 나타내었다.

표 3은 학습에 70%, 테스트에 30%를 사용하였으며, 표 4는 학습과 테스트에 각각 50%의 데이터를

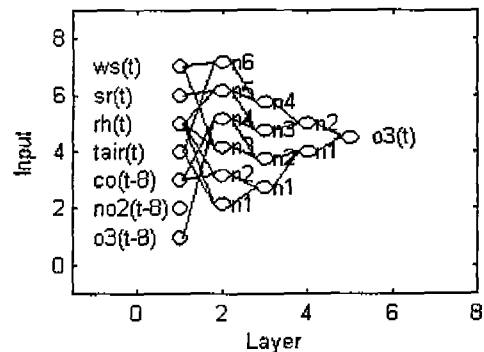


그림 10. 쌍문동 저농도 모델의 구조
Fig. 10. Network structure of Ssang-Mun-Dong low ozone model

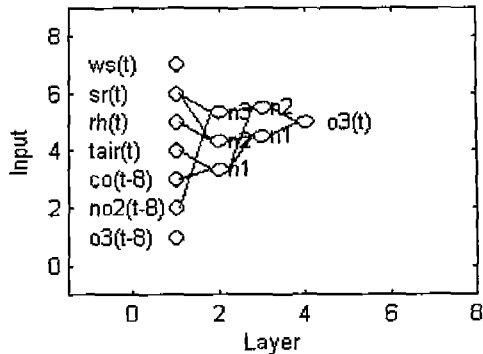


그림 11. 쌍문동 고농도 모델의 구조

Fig. 11. Network structure of Ssang-Mun-Dong high ozone model

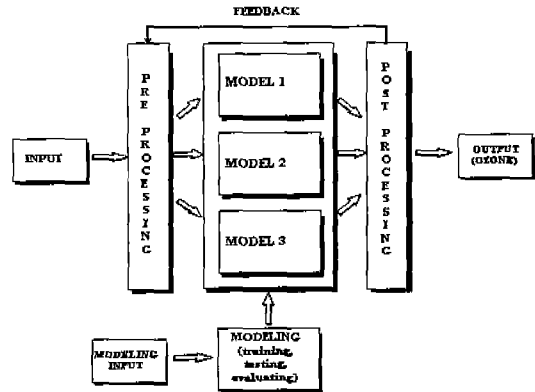


그림 12. 오존 예측시스템의 구조

Fig. 12. Structure of ozone prediction system

표 3. 데이터 분리가 70%인 경우 선택된 입력 변수
Table 3. Selected variables with data split(70%)

구분	모델	선택된 입력 변수	층수
쌍문동	저농도	오존, 일산화탄소, 기온, 습도, 일사량, 풍속	4층
	고농도	일산화탄소, 기온, 습도, 일사량, 풍속	3층
	전체농도	일산화탄소, 기온, 습도, 일사량, 풍속	3층

표 4. 데이터 분리가 50%인 경우 선택된 입력 변수
Table 4. Selected variables with data split(50%)

구분	모델	선택된 입력 변수	층수
쌍문동	저농도	오존, 일산화탄소, 기온, 습도, 일사량, 풍속	3층
	고농도	이산화질소, 일산화탄소, 기온, 습도, 일사량	3층
	전체농도	이산화질소, 일산화탄소, 기온, 습도, 일사량	3층

1996~97. sangall : scatter plot

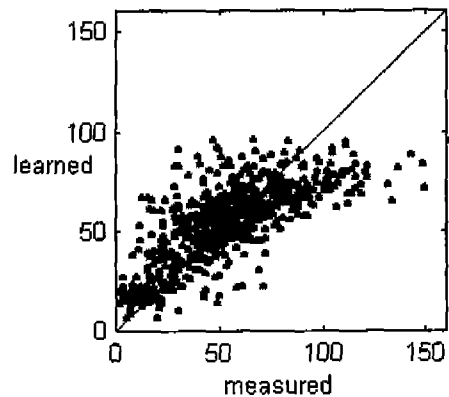


그림 13. 쌍문동 전체농도의 학습결과

Fig. 13. Training result of Ssang-Mun-Dong

사용한 경우이다. 기온과 습도, 일사량은 모든 모델에서 입력변수로 선택되었으며, 대체적으로 데이터 분리가 70%인 경우는 일산화탄소와 풍속, 데이터 분리가 50%인 경우에는 이산화질소와 일산화탄소 등이 추가적으로 오존 농도에 영향을 미치는 변수임을 알 수 있다. 그리고 모든 모델을 통해서 오존이 가장 적은 빈도로 선택되었는데, 이는 단 시간 내에 저농도에서 고농도로 변하는 오존의 특성으로 인해 1~2 시간전의 수치가 아닌 8시간전의 오존 수치는 8시간 후의 오존농도에 크게 영향을 미치지 못하는 것으로 볼 수 있다.

그림 12에 오존 예측 시스템의 전체적인 구조를 나타내었다. 모델링 입력을 통해서 세 가지의 모델(저농도 모델, 고농도 모델, 전체농도 모델)과 각 모델에서 어떠한 변수들이 오존 농도에 영향을 미치는지를 평가한 다음 성능이 우수한 모델을 해당 농도의 예측

모델로 정한 후 실질적인 예측모델로 사용한다.

내일의 오존 농도를 예측하기 위해서 예측 입력으로 들어온 자료를 분석하여(기온, 습도 등) 세 가지의 모델 중 어느 쪽의 모델을 사용해서 오존의 농도를 예측할 것인가를 전처리 부분에서 결정하게 된다. 전처리 부분에서 정해진 모델을 이용하여 오존농도를 예측한 다음, 후처리 부분에서 예측된 오존농도가 적절한 한지를 평가하게 된다. 만약 전처리 부분에서 어떠한 모델로 예측을 해야할지 결정하기 어려운 경우에는 세 가지의 모델로 모두 출력을 구한 후, 후처리 부분에서 각 출력 값에 대하여 가중치를 곱하고 이 값들을 더해서 출력을 구한다. 또한 입력자료를 바탕으로 고농도 모델로 예측을 하였더라도, 오존 농도에 큰 영향을 미치는 강수량이 모델링 과정에서 입력변수로 안 들어갔기 때문에, 전날부터 예측시간 전까지의 강수량이 있을 경우는 후처리 부분에서 다시 저농도 모델이나

전체 농도 모델로 예측한다.

그림 13~16에 쌍문동에 대한 학습결과와 예측결과를 나타내었다.

그림 13은 쌍문동의 전체 자료를 이용하여 학습한 결과이며, 그림 14는 쌍문동 전체농도에 대하여 학습된 모델을 이용해서 97년 8, 9월 달의 쌍문동의 14시의 오존농도를 예측한 결과이다. 오차는 각각 18.79와 18.89이다.

그림 15와 16에 쌍문동 고농도 자료에 대한 학습결과와 97년 8, 9월 14시에 대한 고농도 오존 예측결과를 나타내었다. 오차는 각각 11.65와 17.33이며, 각 모델의 오차를 표 5에 나타내었다.

마지막으로 새로운 자료를 바탕으로 모델을 업데이트시켰을 때의 결과를 아래에 나타내었다. 하루마다

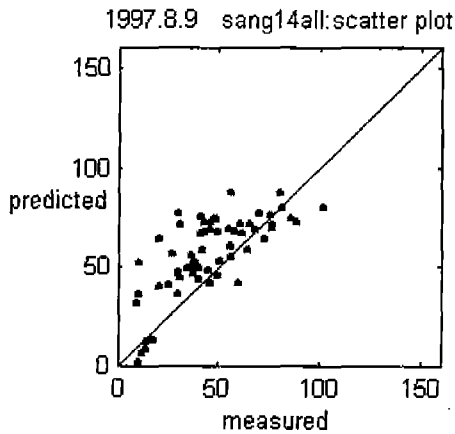


그림 14. 쌍문동 97년 8~9월 전체농도 예측결과
Fig. 14. Prediction result of Ssang-Mun-Dong (1997. 8~9.)

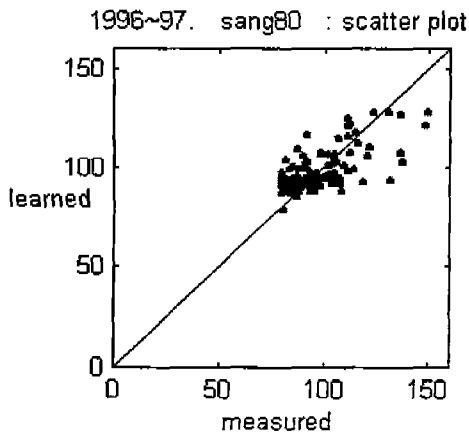


그림 15. 쌍문동 고농도 자료에 의한 학습결과
Fig. 15. Training result of high ozone concentration in Ssang-Mun-Dong

업데이트하는 경우와 일주일에 한번씩 업데이트하는 경우를 고려하였고, 업데이트 된 자료만큼 과거의 자료를 삭제하였다. 쌍문동 97년 8월 1일부터 8월 14일까지 2주 동안의 14시 오존 농도의 예측 결과를 업데이트하지 않은 모델과 비교하였다. 그림 17은 하루씩 업데이트 된 모델의 결과이고, 그림 18은 일주일씩 업데이트 된 모델이며, 그림 19는 업데이트하지 않은 모

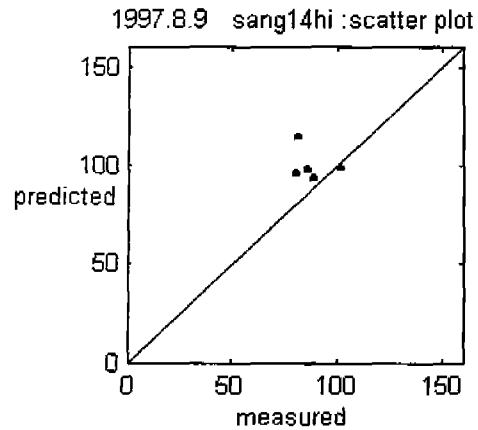


그림 16. 쌍문동 97년 8~9월 고농도 예측결과
Fig. 16. Prediction result of high ozone concentration in Ssang-Mun-Dong (1997. 8~9.)

표 5. 각 모델의 오차
Table 5. Error of each model

구 분	모 델	학습오차	예측오차(14시예측)
	저 농도	13.22	14.61
쌍문동	고 농도	11.65	17.33
	전체농도	18.79	18.89

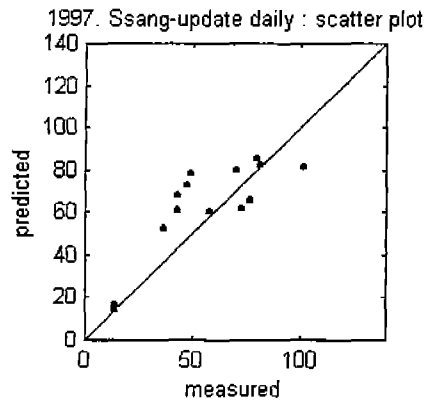


그림 17. 하루씩 업데이트 한 모델
Fig. 17. Model updated daily

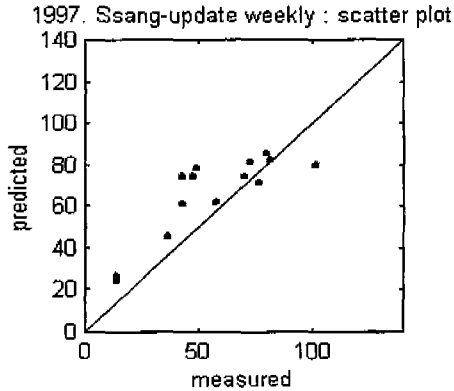


그림 18. 일주일씩 업데이트 한 모델
Fig. 18. Model updated weekly

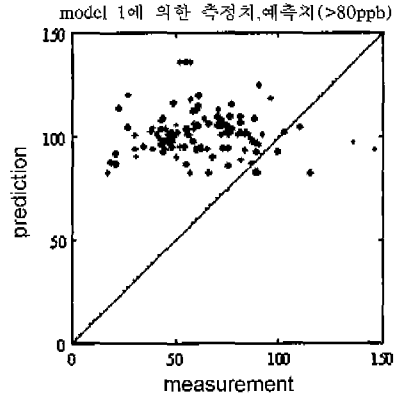


그림 20. 측정치와 80 ppb이상의 예측치 비교
Fig. 20. Measurement and Predictions over 80 ppb

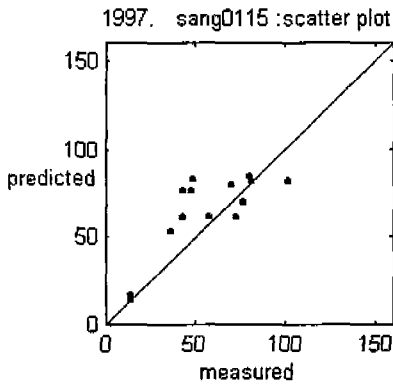


그림 19. 업데이트되지 않은 모델
Fig. 19. Model without updating

델의 결과이다.

오차는 각각 16.61, 15.69, 17.59이며, 일주일씩 업데이트 된 모델의 오차가 가장 작게 나타났다.

4.2 퍼지 예측기

퍼지논리 시스템에 의한 예측기는 동적 다항식 신경회로망의 경우와 마찬가지로 아래와 같은 쌍문동에서의 오존농도를 학습 및 예측 대상으로 적용하였다. 학습데이터는 고농도 오존의 발생이 비교적 빈번했던 96년 5월부터 9월, 97년 6월부터 7월까지의 데이터를 사용하였고, 특별히 오존 농도가 높게 나타나는 14시부터 17시까지 한 시간 간격의 데이터를 사용하였다. 예측대상은 97년 8월부터 9월까지 두 달 동안 14시부터 17시까지의 오존농도이다.

퍼지 예측기를 이용하여 두 가지의 실험을 시행하였다. 첫 번째 실험은 위에서 언급한 학습 및 예측 대상 중 고농도의 오존 경보가 요구되는 80 ppb 이상

의 데이터만을 대상으로 한 경우이다. 이 때 학습에 사용된 데이터는 159개로 전체 데이터 980개 중 16.2%에 해당된다.

실제로 80 ppb 이상의 측정치를 기록한 데이터의 수도 전체 244개 중 11.5%에 해당하는 28개이었다. 다른 실험은 일반적인 경우로 사용할 수 있는 모든 학습 데이터를 사용하여 예측 구간의 14시부터 17시 사이의 오존 농도를 예측하였다. 이상의 두 경우 모두 예측치를 발생하는 시간은 예측 시간을 기준으로 8시간 이전이다. 즉 6시에 14시의 예측치를 발생한다.

A. {실험 1: 80 ppb 이상 오존 농도의 예측}

학습데이터: 쌍문동의 96년 5월부터 9월, 97년 6월부터 7월까지의 (14시부터 17시까지 시간별 데이터) 오존 농도 중 80 ppb 이상의 데이터

예측: 97년 8월과 9월 (14시부터 17시까지 시간별 오존농도)

그림 20은 위의 학습데이터를 사용하여 학습한 후, 80 ppb 이상으로 예측된 결과와 실제치를 보여주며, 그림 21은 예측치와 실제치가 모두 80 ppb 이상으로 나온 결과를 나타낸다.

B. {실험 2: 전범위 오존 농도 예측}

학습데이터: 쌍문동의 96년 5월부터 9월, 97년 6월부터 7월까지의 (14시부터 17시까지 시간별 데이터) 모든 오존 농도 데이터

예측: 97년 8월과 9월 (14시부터 17시까지 시간별 오존농도)

이상에서 살펴 본 바와 같이 실험 1과 실험 2는 학습

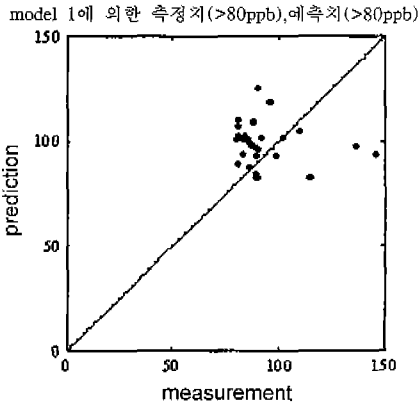


그림 21. 80 ppb 이상의 측정치와 80 ppb이상의 예측치 비교
Fig. 21. Measurement over 80 ppb and Predictions over 80 ppb

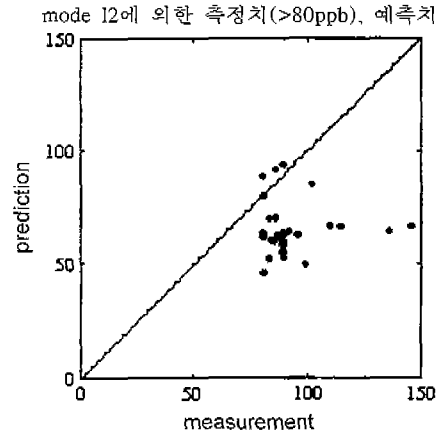


그림 23. 80 ppb 이상의 측정치와 예측치 비교
Fig. 23. Comparison of measurements with predictions over 80 ppb

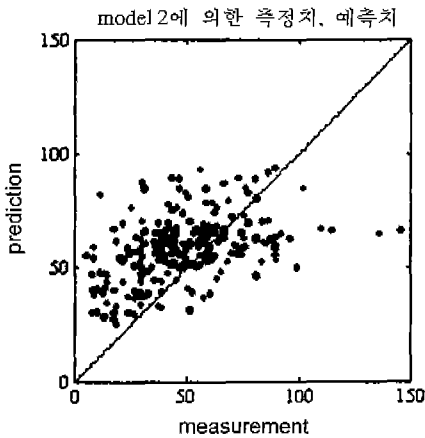


그림 22. 측정치와 예측치 비교
Fig. 22. Comparison of measurements with predictions

에 사용된 데이터가 다르다는 점에서 그 성능의 차이를 보여주고 있다. 그림 20은 80 ppb 이상의 데이터를 이용하였기 때문에 80ppb 이상이라고 판단되는 예측치와 실제의 측정치를 비교하였다. 이에 비해 그림 21은 그림 20의 일부로서 80 ppb 이상이라고 보여준 측정치에 해당하는 예측치를 보여주고 있다. 두 경우의 RMSE값은 각각 45.67와 20.36이었다. 그림 22와 그림 23을 얻기 위해서는 모든 데이터를 사용하여 학습하였다. 그림 22는 모든 값에 대해 예측을 시도한 것이며 그림 23은 80 ppb 이상의 측정치에 대해 예측치를 비교한 것이다. 이 두 경우의 RMSE값은 각각 23.98과 33.73을 얻었다.

위의 그림 24는 14시에서의 측정치, 80 ppb 이상으로 된 학습데이터에 의한 예측치, 그리고 모든 학

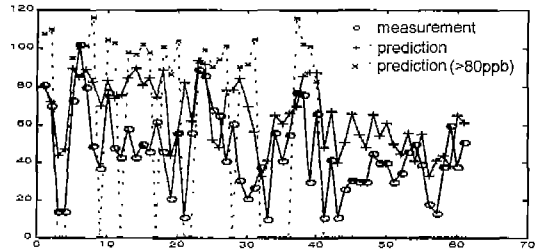


그림 24. 14시에서의 측정치 : 실험 1의 예측치(x), 실험 2의 예측치(+)
Fig. 24. Measurement at 14:00: prediction of the experiment 1 (x), prediction of the experiment 2 (+)

습데이터를 사용하여 얻은 예측치를 보여주고 있다. 특히 80 ppb 이상에 의한 예측치에서는 농도가 0으로 판정된 경우가 있는데 이는 입력에 따른 규칙이 존재하지 않는 것으로 고농도가 아닌 경우를 표현하고 있다.

이상에서 살펴본 바와 같이 RMSE와 같은 어려값이 가지는 의미는 특정 집단에 대한 측정치를 얼마나 잘 예측할 수 있는가 하는 평가기준으로 사용할 수 있다. 그러나 농도 오존의 예측에 있어 고농도일 때의 어려가 저농도일 때의 그것보다 훨씬 중요하다. 이를 좀더 정성적으로 분석하면 측정치가 고농도일 때 예측치를 고농도 혹은 저농도로 예측하는 능력에 더 주목할 필요가 있다. 따라서 아래와 같은 4가지의 경우에 대한 조건부 확률을 구하였다.

실험 1에서 고농도 측정치가 나왔을 때 고농도 예측한 확률: $P(\text{예측치} > 80 | \text{측정치} > 80) = 1$.

실험 1에서 저농도 측정치가 나왔을 때 고농도 예측한 확률: $P(\text{예측치} > 80 | \text{측정치} < 80) = 0.38$.

실험 2에서 고농도 측정치가 나왔을 때 고농도 예측한 확률: $P(\text{예측치} > 80 | \text{측정치} > 80) = 0.18$.

실험 2에서 저농도 측정치가 나왔을 때 고농도 예측한 확률: $P(\text{예측치} > 80 | \text{측정치} < 80) = 0.065$.

이상에서 실험 1에서 사용한 모델을 이용하면 측정치가 고농도일 경우, 예측치도 반드시 고농도가 된다. 그러나 측정치가 저농도임에도 불구하고 예측이 고농도로 나온 오경보율 (false alarm rate)도 38%나 된다. 따라서 실험 1에서 사용한 모델을 이용하면 비록 오경보율이 높더라도 고농도를 저농도로 예측하는 경우는 없다고 단정할 수 있다.

실험 2에서 사용한 모델은 고농도의 측정치에 대해 고농도의 예측한 비율이 18%에 지나지 않는다. 이는 학습에 사용한 데이터가 고농도 예측에는 적절하지 못함을 보여주고 있다. 그러나 저농도의 측정치에 대해 고농도의 예측치를 발생하는 오경보율을 실험 1의 모델에 비해 월등히 적은 6.5%에 지나지 않는다. 따라서 실험 2의 모델은 저농도 예측에 적절한 모델이라고 할 수 있다. 실제로 측정치가 저농도일 때 예측치도 저농도로 나오는 비율은 93.5%가 된다.

5. 결 론

본 논문에서 오존 농도의 발생 모델이 가지고 있는 비선형성을 기존의 해석적인 방법보다 더 잘 학습할 수 있도록 하기 위하여, 동적 다항식 신경회로망 및 퍼지 논리 시스템을 이용하였고, 현실적으로 모든 입력변수를 사용하는 것이 부적당하기 때문에 동적 다항식 신경회로망을 이용하여 입력변수의 수를 줄임으로써, 향상된 예측 성능을 보여주었다.

대부분의 시뮬레이션 결과 동적 다항식 신경회로망이 7개의 입력변수 중에서 선택하는 입력변수는 이산화질소, 일산화탄소, 기온, 습도, 일사량 등이다. 따라서 다른 입력변수인 오존과 풍속 등은 실제 오존 농도 예측에 큰 영향을 주지 않는 것으로 나타났으며, 최적의 성능을 나타내는 모델을 구하기 위해 PC를 도입하여 과도학습을 방지하였다. 또한, 오존 농도 예측에 있어서 좀 더 정확한 예측을 위해서는 모델의 구조를 고정시키는 것보다는 매일의 새로운 자료를 모델의 새로운 입력변수로 사용하여서 예측모델의 구조를 조정해 나가는 것이 더 바람직한 예측의 형태가 될 것으로 생각된다.

퍼지 논리에 의한 예측기의 결과는 동적 다항식 신

경회로망을 사용한 경우와 비교해 볼 때 부정확하지 만, 고농도 예측을 위해서는 고농도로 학습한 예측기가 전체농도를 학습한 예측기 보다 우수한 성능을 보이며, 특히 측정치가 고농도(80 ppb이상)일 때는 예측치가 반드시 고농도임을 확인할 수 있다.

전체적으로 고농도 부분에서 오차가 크게 발생하였으며, 실제로 고농도 데이터의 양이 부족했기 때문에 충분한 학습이 이루어지지 않은 것으로 생각할 수 있다.

향후 과제로서 동적 다항식 신경회로망의 경우 노드 출력 다항식의 형태를 다양화하여 그 성능을 개선시킬 필요가 있으며, 퍼지 예측기의 경우 퍼지 멤버십 함수의 지지영역변경으로 최적의 멤버십 함수를 구하는 것이 필요하다. 마지막으로 더 많은 데이터를 이용한 학습과 오존 농도가 가진 비정상적인 경향을 반영할 수 있는 추가적인 입력이 요구된다.

감사의 글

본 연구는 1998년도 학술진흥재단 학제간 공동연구 지원과제인 「오존예보시스템의 개발」의 일환으로 수행되었으며, 지원에 감사드립니다.

참고문헌

- [1] A. G. Ivahnenko, "Polynomial theory of complex system," *IEEE trans. Syst. Man and Cybern.*, vol. SMC-12, pp. 364-378, 1971.
- [2] 김용국, 이종범, "하계의 일 최고 오존농도 예측을 위한 신경망 모델의 개발," *한국 대기보전학회지*, Vol. 10, No. 4, pp. 224-232, 1994.
- [3] 김영성, "1991~1993년 사이 우리 나라 오존 농도 변화," *한국 대기보전학회지*, Vol. 12, No. 1, pp. 55-66, 1996.
- [4] Duc Trung Pham and Liu Xing, *Neural Networks for Identification, Prediction and control*, Springer-Verlag Inc., 1995.
- [5] S. Chen and S. A. Bilings, "Neural networks for nonlinear dynamic system modeling and identification," in *Advances intelligent Control*(C. J. Jarris, ed.), ch. 4, Taylor and Francis, 1994.
- [6] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural network," *IEEE Transactions on Neural Networks*, Vol. 1, No. 1, pp. 4-27, 1991.
- [7] 허정숙, 김동술, "다변량 통계분석을 이용한 서울시 고농도 오존의 예측에 관한 연구", *한국 대기보전학회지*, Vol. 9, No. 2, 1993.
- [8] G. E. P. Box and G. M. Jenkins, "Time Series Analysis, Forecasting and Control", *San Francisco, Holden Day*, 1976.

- [9] 이선기, 이종범, "서울지역의 광화학 오염농도추정을 위한 증회귀 모델의 개발," 기상연구논문집, Vol. 8, No. 1, pp. 71-85, 1991.
- [10] 김성신, "A Neuro-Fuzzy Approach to Integration and Control of Industrial Processes : Part I," 퍼지 및 지능시스템 학회 논문지, Vol. 8, No. 6, pp. 58-69, 1998.
- [11] A. G. Ivakhnenko, "The Group Method of Data Handling in Prediction Problem," *Soviet Automatic Control*, Vol. 9, No. 6, pp. 21-30, 1976.
- [12] J. Connors, D. Martin, and L. Atlas, "Recurrent neural networks and robust time series prediction," *IEEETrans. Neural Network*, Vol. 5, pp. 240-254, Mar. 1994.
- [13] 김용국, "수도권 지역의 광화학 오존농도 예측에 관한 연구," 강원대 박사학위 논문, 1996.
- [14] 홍낙기, 이종범, 김용국, "다변량 해석을 이용한 서울시 하계스모그의 형태 분류," 한국 대기보전학회지, Vol. 9, No. 4, pp. 278-287, 1993.
- [15] J. S. R. Jang, C. T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing*, Prentice-Hall Inc., 1997.
- [16] S. Farlow, ed., *Self Organizing Method in Modeling: GMDH-Type Algorithms*, Marcel Dekker, Inc., New York, 1984.
- [17] 허정숙, 김동술, "군집분석을 이용한 고농도 오존생성의 사례 연구," 한국 대기보전학회 춘계학술대회 요지집, Vol. G-1, pp. 127-130, 1998.
- [18] L. X. Wang, Fuzzy systems are universal approximators, *Proc. IEEE International Conf. On Fuzzy Systems*, [1] pp. 1163-1170, 1992.
- [19] F. Scarselli and A. C. Tsoi, Universal Approximation Using Feedforward Neural Networks A Survey of Some Existing Methods, and Some New Results, *Neural Networks*, Vol. 11, No. 1, pp. 15-37, 1998.
- [20] S. H. Lee, I. Kim, "Time Series analysis using fuzzy learning," in *Proc. Int. Conf. Neural Inform. Processing*, Seoul, Korea, Oct. 1994, Vol. 6, pp. 1577-1582.
- [21] L. X. Wang, *Adaptive Fuzzy Systems and Control: Design and Stability Analysis*, Prentice-Hall, 1994.



김태현 (Tae-Hun Kim)

1990년 : 부산대학교 전기공학과 공학사
2000년 : 부산대학교 전기공학과 공학 석사



김성신 (Sung-Shin Kim)

1984년 : 연세대학교 전기공학과 공학사
1986년 : 연세대학교 전기공학과 공학 석사
1996년 : Georgia Inst. of Tech. 공학박사
1996년~1998년 : Appalachian Electronic Instruments Inc. 연구원
1998년~현재 : 부산대학교 전기공학과 조교수



김인택 (In-Taek Kim)

1984년 : 서울대학교 전자공학과 학사
1986년 : 서울대학교 전자공학과 석사
1992년 : Georgia Institute of Technology, Electrical Engineering, 박사
1993년~1994년 : LG 전자 기술원
1995년 이후 : 명지대학교



이종범 (Chong-Bum Lee)

1973년 : 서울대(이학사-기상학)
1976년 : 서울대(이학석사-미기상학)
1985년 : 일본 쓰쿠바대(이학박사-대기 환경학)
1989년~1991년 : 강원대 부설 환경연구소 소장
1993년~1994년 : 미국 환경보호청 객원 교수

1995년~현재 : 환경부 중앙환경심의위원회 위원
1998년~현재 : 한국 대기환경학회 부회장



김신도 (Shin-Do Kim)

1976년 : 한양대학교 건축공학과
1978년 : 한양대학교 건축환경 석사
1983년 : 동경공업대학교 Environmental Engineering 박사
1984년~현재 : 서울시립대 환경공학부 교수



김용국 (Yong-Goog Kim)

1986년 : 강원대학교 환경학과 이학사
1992년 : 강원대학교 환경학과 이학석사
1996년 : 강원대학교 환경학과 이학박사
1995년~2000년 7월 : 강원대학교 환경연구소 객원연구원
2000년 : 국립환경연구원 환경기술 진흥센터 전문위원